Research Statement

Minerva Catral

My general area of research is linear and multilinear algebra, with emphasis on the theory of matrices and its applications. I have been involved in several research projects, three of which have successfully culminated in papers that have either been published or accepted for publication.

A class of matrices that arise frequently in applications is the class of M–matrices. These are matrices of the form $sI - B$, where B is (entry-wise) nonnegative and $\rho(B) \leq s$. The book by Berman and Plemmons [2] lists more than fifty equivalent conditions for a matrix of this form to be a non-singular M–matrix. For instance, *A is a nonsingular M–matrix if and only if A^{-1} is a nonnegative matrix.* This makes the study of M–matrices inseparable from the study of nonnegative matrices which is a very rich theory.

My Ph.D. thesis explores some applications of the theory of M–matrices and the theory of nonnegative matrices to finite Markov chains. In particular, I studied finite ergodic Markov chains and their associated mean first passage matrices. In his excellent 1975 paper [17], Meyer has provided expressions for the mean first passage matrix of an n–state homogeneous ergodic Markov chain with states $\{\mathcal{S}_1, \cdots, \mathcal{S}_n\}$, in terms of the group inverse $A^{\#}$ of the M–matrix $A = I - T$, where T is the transition matrix of the chain. On the other hand, the paper by Dietzenbacher [8] (but the papers by Cho and Meyer [5] and Higham [9] also point to other references) gives a formula for the mean first passage time from state $\mathcal{S}_i$ to state $\mathcal{S}_j$ in terms of the row sums of the leading $(n-1) \times (n-1)$ principal submatrix of A obtained by deleting its j–th row and column.

An interesting quantity that arises in applications is the so-called *Kemeny constant.* If T is the transition matrix of an n–state homogeneous ergodic Markov chain with stationary distribution vector $\pi = [\pi_1, \cdots, \pi_n]^T$, and with mean first passage matrix $M = (m_{i,j})$, then it is known that the expressions

$$\sum_{j=1}^{n} m_{i,j}\pi_j, \quad i = 1, \cdots, n$$

all have a common value K, called the Kemeny constant (associated with T). Furthermore, it can be shown that $K - 1$ is equal to the trace of the group inverse $A^{\#}$, where $A = I - T$. Practical interpretations for this constant have been given by Hunter [12] and Levene and Loizou [15]. I have considered the effects of different types of perturbations on T to its Kemeny constant. For instance, if we start with a symmetric T and if we let $\bar{T} = T + E$, where E is positive semidefinite and $\bar{T}$ is still stochastic and irreducible, then $K \leq \bar{K}$. For the general case, I have found a new formula for K in terms of the trace of the inverse of any $(n-1) \times (n-1)$ principal submatrix of A and any of the ratios $\dfrac{A^{\#}_{j,j}}{\pi_j}$, $j = 1, \cdots, n$.

Another topic that can be studied in the context of group inverses and mean first passage matrices is the convexity (or concavity) of the Perron root $\rho(S)$ of a nonnegative irreducible matrix S. In particular, Deutsch and Neumann [6] have given representations for the matrix of second-order derivatives of the Perron root $\rho(S)$ in terms of the group inverse $A^{\#}$, where $A = \rho(S)I - S$. Their work imply that the convexity of the Perron root at the (i,j) entry $(i \neq j)$ is determined by the sign of the (j,i) entry of $A^{\#}$. An open problem was posed by the authors in [6]:

($\star$) *Identify the set of all $n \times n$ nonnegative irreducible matrices S such that the Perron root at S is a concave function at every off-diagonal position.*

The problem above is equivalent to characterizing all the nonnegative irreducible matrices S such that the group inverse $(\rho(S)I - S)^{\#}$ is an M–matrix. Using the relations by Meyer and Dietzenbacher, we were able to give necessary and sufficient conditions for ($\star$) in terms of the mean first passage times, and also in terms of the row sums of the various $(n-1) \times (n-1)$ principal submatrices of A.

A more recent application of the theory is on ring networks that have so-called "small-world" properties. The paper by Watts and Strogatz [20] describe these networks as highly clustered, like regular lattices, yet have small characteristic path lengths, like random graphs. Such networks can be studied analogously using matrix analysis and Markov chain theory. This was done for instance by Higham in his paper [9] where he considered a basic ring network consisting of n vertices where each vertex is connected to its neighboring vertices giving rise to a transition matrix P. For a fixed vertex j, the *maximum mean hitting time* is defined by

$$\text{mht}_{\max} = \max_{1 \leq i \leq n-1} (\bar{M}_j)_i, \tag{0.1}$$

and the *average mean hitting time* is defined by

$$\text{mht}_{\text{ave}} = \frac{1}{n-1} \sum_{i=1}^{n-1} (\bar{M}_j)_i, \tag{0.2}$$

where $\bar{M}_j = [m_{1,j}, \cdots, m_{j-1,j}, m_{j+1,j}, \cdots, m_{n,j}]^T$. (That is, $\bar{M}_j$ gives the mean first passage times $m_{i,j}$ of going into vertex j from any vertex i, $i \neq j$.)

Next, the network is modified by adding "short-cuts", more precisely, each vertex is connected to all its non-neighboring vertices with equal probability ε, and the new matrix that arises, $P(\varepsilon)$ is just a perturbation of P. Higham looks at the ratios

$$\frac{\mathrm{mht}_{\max}(\varepsilon)}{\mathrm{mht}_{\max}(0)}, \tag{0.3}$$

and

$$\frac{\mathrm{mht}_{\mathrm{ave}}(\varepsilon)}{\mathrm{mht}_{\mathrm{ave}}(0)} \tag{0.4}$$

to study how the maximum mean hitting times and the average mean hitting times behave as the parameter ε is increased from 0. The expression in (0.1) is just the maximum row sum of the inverse of the principal submatrix A_j of A, where $A = I - P$ and A_j is obtained by deleting the j–th row and column of A. On the other hand, the expression in (0.2) is an average of the entries in $\bar{M}_j$. For my thesis, I instead look at the expression

$$\sum_{i \neq j} \pi_i m_{i,j}, \tag{0.5}$$

where $\pi = [\pi_1, \cdots, \pi_n]^T$ is the left Perron vector of P. Note that (0.5) is a quantity that is close to the weighted average of the entries in $\bar{M}_j$. As the numbers π_i, $i = 1, \cdots, n$, give the long-run fraction of time that the process is in vertex i, then looking at (0.5) instead of (0.2) seems more reasonable. There is also a matrix-theoretic reason for considering the expression given in (0.5). It can be shown, using Meyer's results (see [17]) that the expression in (0.5) is equal to $\frac{A^{\#}_{j,j}}{\pi_j}$. Hence, the problem of investigating the ratios in (0.4) becomes a perturbation problem on the ratios of the j-th diagonal entry of $A^{\#}$ and the j-th entry of the left Perron vector π of P.

Research Projects

- **On Functions that Preserve M–matrices and Inverse M–Matrices** (with R. Bapat and M. Neumann), *Linear and Multilinear Algebra*, to appear. This work surveys some results on the question of functions that preserve the entire class or a sub-class of the M–matrices or that map them to the nonnegative matrices. The main results of this paper are summarized by three theorems that characterize functions which either preserve the inverse M–matrices (functions whose inverses are M–matrices) or map the inverse M–matrices to the M–matrices.

- **On Reduced Rank Nonnegative Matrix Factorization for Symmetric Nonnegative Matrices** (with L. Han, M. Neumann and R. Plemmons), *Linear Algebra and Its Applications*, 393:107–126, 2004. Let $V \in \mathbb{R}^{m,n}$ be a nonnegative matrix. The *nonnegative matrix factorization* (NNMF) problem consists of finding nonnegative matrix factors $W \in \mathbb{R}^{m,r}$ and $H \in \mathbb{R}^{r,n}$ such that $V \approx WH$. We have considered the algorithm provided by Lee and Seung [13, 14] which finds nonnegative W and H such that $\|V - WH\|_F$

is minimized. For the case $r = 1$, we have given a complete characterization of the solution. We next considered the case $m = n$ and in which V is symmetric. In this work, we focused on questions concerning when the best approximate factorization results in the product WH being symmetric and on cases in which the best approximation cannot be a symmetric matrix. We have also shown that any positive semidefinite symmetric nonnegative matrix V generated by a Soules basis admits, for every $1 \leq r \leq \text{rank}(V)$, a nonnegative factorization WH which coincides with the best approximation in the Frobenius norm to V in $\mathbb{R}^{n,n}$ of rank not exceeding r.

- **Proximity in Group Inverses of M–Matrices and Inverses of Diagonally Dominant M–Matrices** (with M. Neumann and J. Xu), *Linear Algebra and Its Applications*, to appear.
  In this recent work we have improved upon and generalized results concerning Laplacians of undirected weighted graphs and inverses of strictly diagonally dominant M–matrices. For instance, if L is the Laplacian matrix of a connected graph $\mathcal{G}$ consisting of n vertices, then Chebotarev and Shamis [4] showed that if $Q = (I + L)^{-1} = (q_{i,j})$, then the following proximity relations hold for any triple of indices $1 \leq i, j, k \leq n$:

$$q_{i,i} - q_{i,j} - q_{i,k} + q_{j,k} \ \geq \ 0. \tag{0.6}$$

  We have generalized this result for any matrix of the form $A = D - T$, where T is nonnegative, irreducible and symmetric, and D is the diagonal matrix whose i-th diagonal entry is the i-th row sum of the matrix T. Note that A in this case is singular but it possesses a group inverse $A^{\#}$. One of our results is that the relation in (0.6) holds when Q^{-1} is replaced by $A^{\#}$.

Research Plans

In the Fall of 2003, my adviser has directed me to some papers [18, 19] in physical chemistry, particularly on calculating reaction paths in the process of photosynthesis using mean first passage times. The authors describe the "excitation migration" in PSI (Photosystem I), a protein pigment consisting of 96 chlorophylls, including the special pair P700. This process can be viewed as a continuous-time, discrete-state ($n = 96$) Markov chain, with the two chlorophylls in P700 being absorbing states as there is charge separation when any of the 94 chlorophylls reach any of these two states. Notice that the resulting system is an absorbing Markov chain, and hence the techniques employed for finite ergodic Markov chains described earlier do not directly apply here.

While looking at these papers, we noticed that the equations that arise involve nonsingular, diagonally dominant M–matrices. For instance, they define the intensity matrix $K = (K_{i,j})$ by

$$K_{i,j} = \begin{cases} R_{i,j}, & i \neq j, \\ -\sum_{\ell \neq i} R_{i,\ell} - \lambda_i, & i = j, \end{cases}$$

where $R_{i,j}$ is the transfer rate from state i to state j, and λ_i is the probability loss at state i. We have considered the problem of determining the effects on the probability vector ϕ (whose entries ϕ_i give the probabilities that passage to the product set will occur if started at state i) and on the mean first passage vector τ (whose entries τ_i give the mean first passage times from state i to the product set), if a first-order perturbation is applied to the intensity matrix K. We had some small results along these lines, but have not had time to fully explore the problem. This will be an interesting topic for future research.

My long-term research goal is to be able to branch out to other areas of mathematics that are related to matrix theory. The natural areas that I have in mind are in numerical analysis and computational mathematics, and functional analysis and operator theory. If given the opportunity, I would love to be able to do research on the operator theory aspect of matrix theory. I am confident that my research background and exposure to many different areas of study in my current field makes me highly capable of pursuing these research interests.

References

[1] A. Ben–Israel, T.N.E. Greville, Generalized Inverses: Theory and Applications, 2nd ed., Springer, New York, 2003.

[2] A. Berman and R. J. Plemmons. Nonnegative Matrices in the Mathematical Sciences. *SIAM Publications*, Philadelphia, 1994.

[3] S. L. Campbell and C. D. Meyer. Generalized Inverses of Linear Transformations. *Dover Publications*, New York, 1991.

[4] P. Yu. Chebotarev and E.V. Shamis. The Matrix-Forest Theorem and measuring relations in small social groups. *Institute of Control Sciences*, Russian Academy of Sciences, 1997.

[5] G. E. Cho and C. D. Meyer. Markov chain sensitivity measured by mean first passage times. *Lin. Alg. Appl.*, 316:21–28, 2000.

[6] E. Deutsch and M. Neumann. Derivatives of the Perron root at an essentially nonnegative matrix and the group inverse of an M-matrix. *J. Math. Anal. Appl.*, 102:1–29, 1984.

[7] E. Deutsch and M. Neumann. On the first and second order derivatives of the Perron vector. *Lin. Alg. Appl.*, 71:57–76, 1985.

[8] E. Dietzenbacher. Perturbations of the Perron vector: Applications to Finite Markov Chains and Demographic Population Models. *Environment and Planning*, 22 no. 6:747–761, 1990.

[9] Higham, N. A matrix perturbation view of the small world phenomenon. *SIAM J. Mattrix Anal. Appl.*, 25. no. 2:429–444, 2003.

[10] J. G. Kemeny and J. L. Snell. Finite Markov Chains. Springer, New York, 1976.

[11] R.A. Horn, C.R. Johnson. Topics in Matrix Analysis. *Cambridge University Press*, New York, 1991.

[12] J. J. Hunter. Mixing Times with applications to Perturbed Markov chains. *Preprint*, Massey Univ., New Zealand, 2003

[13] D. D. Lee and H. S. Seung, Learning the parts of the objects by non-negative matrix factorization, *Nature,* 401 (1999), 788–791.

[14] D. D. Lee and H. S. Seung, Algorithms for non-negative matrix factorization, *Advances in Neural Information Processing,* 2000.

[15] M. Levene and G. Loizou. Kemeny's constant and the random surfer. *Amer. Math. Monthly*, 109:741–745, 2002.

[16] C. D. Meyer. Matrix Analysis and Applied Linear Algebra. SIAM, Philadelphia, 2000.

[17] C. D. Meyer. The role of the group generalized inverse in the theory of finite Markov chains. SIAM Rev., 17:443–464, 1975.

[18] S. Park and K. Schulten. Reaction Pathways Based on the Gradient of the Mean First-Passage Times. arXiv:physics/0207005v3, 2002.

[19] S. Park, M.K. Sener, D. Lu and K. Schulten. Reaction paths based on mean first passage times. *J. Chem. Phys.*, Vol. 119, No. 3: 1313–1319, 2003.

[20] Watts D.J. and Strogatz S.H. Collective dynamics of "small-world" networks. *Nature*, 393:440–442, 1998.